Best Practices for CONTENTdm and other OAI-PMH compliant repositories: creating sharable metadata

Version 3.0

2/7/2012

Please direct correspondence to:

Geri Bunker Ingram, ingramg@oclc.org, OCLC Digital Collection Services

History and acknowledgements:

Throughout the digital repository landscape, it is increasingly accepted that metadata needs not only to serve the local community but also be suitable for harvesting externally. The challenge is to sustain useful local information while providing context and perspective to both the local and the remote user. Because each metadata standard and each collection management toolset may derive its own 'best practice,' it is incumbent upon each community of practice to provide leadership from its constituents' particular points of view.

Thus, in August 2009, OCLC Digital Collection Services (DCS) convened the CONTENTdm Metadata Working Group (MWG) to create a 'best practices' guideline for our community. Discussions followed presentations given at regional and national CONTENTdm Users Groups, and collaborative work was undertaken using the tools familiar to the collective—CONTENTdm, WorldCat Digital Collection Gateway, (Gateway) and various social networking environments. The discussion focused on members' research and publications, and on their efforts to develop, optimize and standardize CONTENTdm metadata element sets such that materials are discoverable easily both in the local CONTENTdm environment as well as across repositories into which their metadata might be harvested according to the standard OAI protocols.

OCLC DCS allocated CONTENTdm servers and trained the MWG members to use the Gateway to map qualified Dublin Core metadata and test them against WorldCat.org displays and WorldCat MARC fields. In the course of the work, the MWG untied several knotty issues and made suggestions resulting in significant improvements to the Gateway. In July, 2010, the Gateway was opened to any OAI-PMH compliant repository.

Table of Contents

Challenges

Essentially there are four types of problems that we see when metadata are viewed outside the context of the collection home. These were generally described in a 2006 article[ii] published by *First Monday*.

Typical problems include:

- Lack of consistency within a single collection.
  -Example: The use of both the Dublin Core <date> and <coverage> elements to record some variant of the resource creation date.

- Too much information.
  -Example: Inclusion of technical information such as date digitized and type of scanner used.

- Lack of key contextual information.
  -Example: Exclusion of a collection name that is essential to make sense of the record.

- Lack of conformance to technical standards.
  -Example: Metadata encoded in XML with character encoding problems.

Recommendations

Likewise, Shreeves (2006) recommends several general practices which CONTENTdm collection administrators would do well to consider. They include:

- We encourage institutions to think carefully about how they might generate *multiple views* of resources using the metadata already created rather than simply sharing a single record describing everything about a resource.

- An institution should understand what an aggregator needs included in the metadata (learning standards? audience level?) to support its service and, when possible, work to meet those needs.

- Metadata aggregators can more effectively normalize records from metadata providers if all records within a defined set are consistent both semantically and syntactically.

- When multiple values are needed, the metadata element should be repeated.

And from M.J. Han, et al, at the University of Illinois[iii] come these further recommendations. Since their research focused on sharing CONTENTdm collection metadata with OAI harvesters, these are especially relevant to our community:

- Keep a balance between specificity and generality in defining local fields.

- Decide at the outset which locally defined fields are intended only for the local environment and which should be made available to aggregators.

- Be cognizant of how values will be created in the local environment.

- Maximize use of Qualified Dublin Core elements for labeling in the local environment.

- Consider taking field names and definitions, if possible, directly from other metadata standards such as EAD, VRA Core, and CDWA when creating locally developed application profiles.

- Share the logic of mapping decisions with aggregators.

Opportunities

In the current metadata aggregation landscape, it is safe to assume that users search and browse for resources at an aggregator's site then follow a link back to the home institution for access to the resource itself and any additional metadata. Therefore, when creating metadata for the purposes of inclusion in these aggregations, one can afford to be selective about the data elements included, with the understanding that a user will find his way to the local records for full contextual information. (Shreeves, 2006)

On July 20, 2009, the OCLC *Digital Collection Gateway* became available to all CONTENTdm 5.1 users in the form of CONTENTdm *WorldCat Sync*. This integrated function enables a CONTENTdm collection administrator to map qualified and simple Dublin Core elements from digital items held in the CONTENTdm collection, to MARC fields, creating and modifying WorldCat records that are synchronized on a schedule set by the collection administrator. The Gateway thus represents a timely opportunity to provide specific Dublin Core metadata schemas for use in CONTENTdm and intended for OAI-PMH harvesting, and underscores a rather urgent need to provide advice to our community.

Below are some notes on creating and configuring metadata for discovery of digital items in WorldCat.org:

- For all fields that you want to display in WorldCat, configure the metadata fields in CONTENTdm so that those fields are mapped to an appropriate Dublin Core element. You can use any Simple Dublin Core and Qualified Dublin Core elements. We recommend using Qualified Dublin Core elements for the best mapping results.

- Date fields should use consistent date formatting.

- Metadata fields set to hidden in CONTENTdm are not available for use with the Digital Collection Gateway.

- If you opt to make a field "Non-Searchable" in CONTENTdm and map that field into the Digital Collection Gateway, the field will be searchable in WorldCat.org.

Core and Recommended Metadata Elements for CONTENTdm Digital Collections and other OAI-PMH compliant sets

*"An element is a descriptive category of information about the resource…. All of the elements used to describe a resource together make up a **record**."- NCSU Libraries Core 1.0 Metadata Element Set Best Practices*

The following is a set of guidelines for understanding using and mapping Dublin Core elements according to the Open Archives Initiative Protocol for Metadata Harvesting. It began as a guide for CONTENTdm collection administrators, and was expanded with the opening of the OCLC Digital Collection Gateway to WorldCat for all OAI-PMH compliant repositories. These guidelines promote the simplification of local information to enable better end-user discovery in an aggregated environment. As with any Best Practices Guide, it is recommended that catalogers follow basic rules of consistency with grammar and syntax (content standard) set forth in resources such as AACR2, DACS, CCO, etc., as well as incorporate the use of controlled vocabularies such as LCSH, AAT, MeSH, and authority lists such as LCNAF and ULAN or 'locally-grown' thesauri as appropriate to the subject matter of a resource. For each digital collection, a collection-level record should be created along with item-level records. Metadata elements should contain labels most useful to the local environment, but should be mapped to standard Dublin Core elements.

A note about repeating fields: A number of works have been published offering best practices for configuring OAI-harvestable metadata. Although these works recommend repeating fields versus multiple values, in some cases multiple values (separated by a semicolon) are preferred for accuracy depending upon the level of complexity in configuring a collection using your digital collections management software and the OAI harvesting tool. For example, semicolon-separated values can be easily accommodated in CONTENTdm as well as display accurately when synced to WorldCat.org via the Digital Collection Gateway. When in doubt, test your data sets against your chosen OAI harvester.

Explanation of Table Components

Element Name	The unique name used in CONTENTdm Version 6.1
DC Definition	Definition as stated in the DCMI Metadata Terms http://dublincore.org/documents/dcmi-terms/
Required DC Element	• *Core, recommended*: the main fields to be used to describe a resource, important for sharing outside of local context. • *Recommended, as appropriate*: the secondary fields which are helpful if available. To use or not depends on the circumstances and the collection manager.
Controlled Vocabulary	Recommended for data quality and consistency
Syntax Scheme	Recommended syntax scheme used to structure the data contained in a given field
DC Element Map	The Dublin Core element to which the CONTENTdm metadata field name maps
MARC Map in WorldCat	The OCLC MARC field to which the Dublin Core metadata element is crosswalked.
Repeatable	• Yes: a field may appear multiple times in a single record • Not preferred: a field should occur only once in a single record
Best Practices	Comments and other recommendations

Core and Recommended Elements

Title

Element Name	Title
DC Definition	A name given to the resource.
Required	Core, recommended
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Title (dc:title)
MARC Map in WorldCat	245
Repeatable	Not preferred

Best Practices
- Prefer literal and non-numeric description of resource, excluding material-type information if possible.
- Prefer non-use of explanatory or qualifying symbols (e.g., brackets to indicate cataloger-supplied title).
- If the recourse has multiple titles (e.g., translated titles, etc.), prefer to use Title-Alternative element.

"Make the title descriptive yet brief. Use generic titles to bring together different images of the same subject, if possible (e.g., use Mayor Benjamin Bosse on all photos of him, so they display together by title)." – Metadata Guidelines, Evansville Photos Collection, Evansville Vanderburgh Public Library.

Title-Alternative

Element Name	Title-Alternative
DC Definition	An alternative name for the resource.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Alternative Title (dcterms:alternative)
MARC Map in WorldCat	246
Repeatable	Yes

Best Practices
- Secondary titles should be used in Title-Alternative.

Creator

Element Name	Creator
DC Definition	An entity primarily responsible for making the resource.
Required	Core, recommended
Controlled Vocabulary	LCNAF, ULAN
Syntax Scheme	
DC Element Map	Creator (dc:creator)
MARC Map in WorldCat	720
Repeatable	Not preferred

Best Practices
- Examples of a Creator include a person, an organization, etc.

- "Prefer use of Name (personal or corporate) Authority Source to be used consistently throughout description of a resource and from one resource to another." - Metadata Implementation Guidelines for North Carolina Digital State Documents
- Prefer non-use of 'junk value' (e.g., "Unknown,") however, it is appropriate to qualify named entities with "[role]".
- WorldCat.org display mapping: dc:creator maps to MARC 720 by default in the Gateway. To enhance precision in fielded searching within WorldCat.org, map dc:creator to MARC 100 (for Personal Name) or 110 (For Corporate Name).

"Do not use honorifics, titles, or nicknames unless it is necessary to disambiguate (e.g., the first name of the person is unknown). Otherwise, these alternate forms of names (such as "Buddy" Jones; Reverend Murrell; Dr. Reed) may be used in the Description field but not as the authoritative version…." – Huntington Digital Library Guidelines, The Huntington Library

Contributors

Element Name	Contributors
DC Definition	An entity responsible for making contributions to the resource.
Required	Core, recommended
Controlled Vocabulary	LCNAF, ULAN
Syntax Scheme	
DC Element Map	Contributor (dc:contributor)
MARC Map in WorldCat	720
Repeatable	Yes

Best Practices
- Examples of a Contributor include a person (e.g., additional writer, illustrator, editor, finding aid author, etc.), an organization, etc.
- Contributors are named so because their responsibility for the creation of a work is not equal to that named as Creator.
- Prefer use of Name (personal or corporate) Authority Source to be used consistently throughout description of a resource and from one resource to another.
- Prefer non-use of 'junk value' (e.g., "Unknown,") however, it is appropriate to qualify named entities with "[role]."

"Persons or organizations who made significant intellectual contributions to the resource, but whose contribution is usually secondary to the person or organization specified in the Creator element. Examples include co-author, editor, transcriber, translator, illustrator, etc." – Metadata Implementation Guidelines for North Carolina Digital State Documents

Description

Element Name	Description
DC Definition	An account of the resource.
Required	Core, recommended
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Description (dc:description)
MARC Map in WorldCat	520 [8]
Repeatable	Yes

Best Practices

- Description may include but is not limited to: an abstract, a table of contents, a graphical representation, or a free-text account of the resource.
- Some digital collections management practitioners prefer the local practice of mapping separate Table of Contents, Abstract, and similar local elements, to Description.
- Prefer collection-based cataloger decision on enabling full-text searching for this field.
  - If data type **Full Text Search**, prefer no mapping to WorldCat. See Description-Abstract element.
  - If data type **Text**, prefer mapping wc.Summary (MARC 520 [8]).

"Also include any other information a searcher might need to find an image through a keyword search or to understand the context of the image: Is there a view of the Mississippi River? Was a photograph taken from the future site of a university library? Does a building no longer exist? What location was a photograph taken from? Is it an aerial view" –WAICU Metadata Guide

Description-Abstract

Element Name	Description-Abstract
DC Definition	A summary of the resource.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Abstract (dcterms:abstract)
MARC Map in WorldCat	520 [3]
Repeatable	Not preferred

Best Practices

- With CONTENTdm, only one Full Text Search field per collection is allowed; therefore if **Description** field is of data type *Full Text Search*, **Description-Abstract** will be of *Text* data type.

Description-Table Of Contents

Element Name	Description-Table Of Contents
DC Definition	A list of subunits of the resource.
Required	Recommended as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Table Of Contents (dcterms:tableOfContents)
MARC Map in WorldCat	505 [8]
Repeatable	Not preferred

Best Practices

- Example:

Chapter 1: Getting Started 1

   Introduction 2

   Next Steps 3

Publisher

Element Name	Publisher
DC Definition	An entity responsible for making the resource available.

Required	Core, recommended
Controlled Vocabulary	LCNAF
Syntax Scheme	
DC Element Map	Publisher (dc:publisher)
MARC Map in WorldCat	260 $b
Repeatable	Yes

Best Practices

- Examples of a Publisher include a person, an organization, etc.
- Prefer use of Name (personal or corporate) Authority Source to be used consistently throughout description of a resource and from one resource to another.
- Prefer non-use of 'junk value' (e.g., "Unknown").
- Prefer "digitized by" or other text prefix to qualify value; Gateway allows both prefix and suffix text constants for each field in every profile.

"The entity responsible for making the Resource available in its present form, such as a corporate publisher, a university department, or a cultural institution." – University of Wisconsin Digital Library Data Dictionary

Subject

Element Name	Subject
DC Definition	The topic of the resource.
Required	Core, recommended
Controlled Vocabulary	DDC, LCC, LCSH, MeSH, UDC, LCNAF, AAT, TGN
Syntax Scheme	
DC Element Map	Subject (dc:subject)
MARC Map in WorldCat	MARC 650 (controlled) / MARC 653 (uncontrolled)
Repeatable	Yes

Best Practices

- Typically, the subject will be represented using keywords, key phrases, or classification codes. Recommended best practice is to use a controlled vocabulary.
- WorldCat.org display mapping: prefer map to MARC 650 if controlled, to MARC 653 if uncontrolled
- To describe the spatial or temporal topic of the resource, use the Coverage element.

"Use subject terms that describe what an object is as well as what it is about. Example 1: Mural painting and decoration; Derry (Northern Ireland); Ireland—History—Easter Rising, 1916." – Guidelines for Metadata Application in the Claremont Colleges Digital Library

Identifier

Element Name	Identifier
DC Definition	An unambiguous reference to the resource within a given context.
Required	Core, recommended
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Identifier (dc:identifier)
MARC Map in WorldCat	856 $u (URL), 024 (non-URL)
Repeatable	Yes*

Best Practices

- *URL*: Gateway selects the first Identifier that contains a URL and makes it the default value for the resolution URL in MARC 856 $u.
    - o If your resolution URL is in a field other than the first Identifier field, you will map it separately.
        - Use the *Edit metadata map* function.
        - Choose the **WorldCat.org** *Item View.*
        - **Click** on the yellow box in the *Find a copy online* section, and map the URL
    - o Thumbnail display images:
        - CONTENTdm supplies the Reference URL to Identifier. This not only provides the resolution URL but also automatically generates the thumbnail for WorldCat.org.
        - OTHER OAI-compliant repositories: To display your thumbnail image in WorldCat.org, with forthcoming Gateway Ver.2.4, select the yellow box labeled Click to map thumbnail URL field under the rectangle anchoring the position for a thumbnail. Then associate one of your source metadata fields with the thumbnail URL.
    - o *Repeatability: It will take all other URLs in repeating Identifier fields, and place them in repeating 856 fields but with no $3 text.

- *Non-URL*: Examples include accession number, ISBN, photo negative job/roll/frame number, call number, etc.
    - o Digital Collection Gateway automatically populates a value for a non-URL Identifier (MARC 024).

"If contributing a digital resource to a collaborative digital collection, consider prefixing the character string with an institutional code to keep your resources distinguishable from those owned by other institutions." –Mountain West Digital Library Metadata Group
Recommended best practice is to identify the resource by means of a string conforming to a formal identification system.

Language

Element Name	Language
DC Definition	A language of the resource.
Required	Core, recommended
Controlled Vocabulary	
Syntax Scheme	ISO 639-2, RFC 1766, RFC 3066, RFC 4646
DC Element Map	Language (dc:language)
MARC Map in WorldCat	546
Repeatable	Yes
Best Practices	

- Multiple values are often used when a resource contains more than one language.

"Separate terms by semi-colon (;) and a space. For example, for French and English: fre; eng" – Metadata Supplement for Fashion Plate Collection, Claremont Colleges Digital Library

Rights

Element Name	Rights
DC Definition	Information about rights held in and over the resource.
Required	Core, recommended
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Rights (dc:rights)
MARC Map in WorldCat	540
Repeatable	Yes

Best Practices
- Prefer free text statement of rights to a 'lonely' URL.
- Rights information includes a statement about various property rights associated with the resource, including intellectual property rights.
- Rights statements should provide references or contact information. Additional clarification can be indicated via linking to an institutional policy statement or other web resource.

"These statements should be given in the form: Rights status. Reproduction/use restrictions. Further information." – Core 1.0 Metadata Element Set Best Practices, NCSU Libraries

Rights-Access Rights

Element Name	Rights-Access Rights
DC Definition	Information about who can access the resource or an indication of its security status.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Access Rights (dcterms:accessRights)
MARC Map in WorldCat	(506##$a enhancement recommended)
Repeatable	Yes

Best Practices
- Access Rights may include information regarding access or restrictions based on privacy, security, or other policies.

Rights-Rights Holder

Element Name	Rights-Rights Holder
DC Definition	A person or organization owning or managing rights over the resource.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Rights Holder (dcterms:rightsHolder)
MARC Map in WorldCat	
Repeatable	Yes

Best Practices
- Prefer to include the name of the copyright holder and the contact information.

Type

Element Name	Type
DC Definition	The nature or genre of the resource.

Required	Core, recommended
Controlled Vocabulary	DCMI
Syntax Scheme	
DC Element Map	Type (dc:type)
MARC Map in WorldCat	655
Repeatable	Yes

Best Practices
- Moving images, three-dimensional objects and sound recordings are all examples of Resource Types.
- Prefer DCMI Type Vocabulary for controlled list of authorized terms: http://dublincore.org/documents/dcmi-type-vocabulary/
- To describe the file format, physical medium, or dimensions of the resource, use the Format element.

"This element should be populated from the DCMI type vocabulary, a controlled listing of genre types. It may be automatically populated, based on characteristics of the repository." – NCSU Libraries Core 1.0 Metadata Element Set Best Practices

Format

Element Name	Format
DC Definition	The file format, physical medium, or dimensions of the resource.
Required	Core, recommended
Controlled Vocabulary	MIME, AAT
Syntax Scheme	
DC Element Map	Format (dc:format)
MARC Map in WorldCat	500 (General Note)
Repeatable	Yes

Best Practices
- Examples of dimensions include size and duration.
- Prefer use of Internet Media Types [MIME] or two-part (type/subtype) identifier in a single string: http://www.iana.org/assignments/media-types/. E.g., audio/mp3; image/jpg; application/pdf; text/html.

"New media types and applications are always emerging. If the resource format being described is not yet part of the MIME type list, select a broad category of object format for the first part of the MIME type, then use the file name suffix for the second half." – University of Louisville CONTENTdm Cookbook

Format-Extent

Element Name	Format-Extent
DC Definition	The size or duration of the resource.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Extent (dcterms:extent)
MARC Map in WorldCat	300
Repeatable	Yes

Best Practices
- Examples include a number of pages, a specification of length, width, and breadth, or a period in hours, minutes, and seconds. E.g., 109,568 bytes; 00:16 minutes.

Format-Medium

Element Name	Format-Medium
DC Definition	The material or physical carrier of the resource.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Medium (dcterms:medium)
MARC Map in WorldCat	300, 340
Repeatable	Yes

Best Practices
- Used for Physical Resource only
- Examples include paper, canvas, or DVD.

Date

Element Name	Date
DC Definition	A point or period of time associated with an event in the lifecycle of the resource.
Required	Core, recommended
Controlled Vocabulary	
Syntax Scheme	W3CDTF
DC Element Map	Date (dc:date)
MARC Map in WorldCat	260 $c
Repeatable	Not preferred

Best Practices
- Date may be used to express temporal information at any level of granularity. Recommended best practice is to use an encoding scheme, such as the W3CDTF profile of ISO 8601 [W3CDTF]. See Appendix B: Dates
- Prefer non-use of 'junk value' (e.g., "Unknown").
- If more than one date is going to be used to describe the resource, it is recommended to use the sub-elements of Date to clarify the type of date, such as Date-Accepted, Date-Issued, etc.
- If only one Date value is present, users may choose to use the Gateway "Prefix/Suffix" feature to explain the context of the date given, e.g., a literal such as "Digitally published". By default, Gateway maps dc:date to MARC 260 $6 (Date of publication, distribution).

"Similarly, if you will describe both physical and digital manifestation properties in your local system using unique field names, consider whether you intend to follow the Dublin Core one-to-one principle, in which case only metadata about one manifestation will be mapped and made available to aggregators." – Metadata for Special Collections in CONTENTdm: How to improve interoperability of Unique Fields through OAI-PMH

Date-Accepted

Element Name	Date-Accepted
DC Definition	Date of acceptance of the resource.

Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	W3CDTF
DC Element Map	Date Accepted (dcterms:dateAccepted)
MARC Map in WorldCat	(502##$a enhancement recommended)
Repeatable	Not preferred

Best Practices
- Examples of resources to which a Date Accepted may be relevant are a thesis (accepted by a university department) or an article (accepted by a journal).
- Prefer to use the "Prefix/Suffix" feature in Gateway with label "Date Accepted". E.g., Date Accepted 2010-03-17.

Date-Submitted

Element Name	Date-Submitted
DC Definition	Date of submission of the resource.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	W3CDTF
DC Element Map	Date Submitted (dcterms:dateSubmitted)
MARC Map in WorldCat	(502##$a enhancement recommended)
Repeatable	Not preferred

Best Practices
- Examples of resources to which a Date Submitted may be relevant are a thesis (submitted to a university department) or an article (submitted to a journal).
- Prefer to use the "Prefix/Suffix" feature in Gateway with label "Date Submitted ". E.g., Date Submitted 2010-03-15.

Date-Created

Element Name	Date-Created
DC Definition	Date of creation of the resource.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	W3CDTF
DC Element Map	Date Created (dcterms:created)
MARC Map in WorldCat	046 $k
Repeatable	Not preferred

Best Practices
- Prefer to use the "Prefix/Suffix" feature in Gateway with label "Date Created".

Date-Available

Element Name	Date-Available
DC Definition	Date (often a range) that the resource became or will become available.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	W3CDTF
DC Element Map	Date Available (dcterms:available)
MARC Map in WorldCat	307 8#

Repeatable	Not preferred
Best Practices	
• Prefer to use the "Prefix/Suffix" feature in Gateway with label "Date Available".	

Date-Valid

Element Name	Date-Valid
DC Definition	Date (often a range) of validity of a resource.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	W3CDTF
DC Element Map	Date Valid (dcterms:valid)
MARC Map in WorldCat	046 $m
Repeatable	Not preferred
Best Practices	
• Prefer to use the "Prefix/Suffix" feature in Gateway with label "Date Valid".	

Date-Copyrighted

Element Name	Date-Copyrighted
DC Definition	Date of copyright.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	W3CDTF
DC Element Map	Date Copyrighted (dcterms:dateCopyrighted)
MARC Map in WorldCat	260 $c
Repeatable	Not preferred
Best Practices	
• Prefer to use the "Prefix/Suffix" feature in Gateway with label "Date Copyrighted ". E.g., Date Copyrighted 2010-03.	
• Both dcterms:dateCopyrighted and dcterms:issued are mapped to MARC 260 $c by default in Gateway.	

Date-Issued

Element Name	Date-Issued
DC Definition	Date of formal issuance (e.g., publication) of the resource.
Required	Core, recommended
Controlled Vocabulary	
Syntax Scheme	W3CDTF
DC Element Map	Date Issued (dcterms:issued)
MARC Map in WorldCat	260 $c
Repeatable	Not preferred
Best Practices	
• Both dcterms:dateCopyrighted and dcterms:issued are mapped to MARC 260 $c by default in Gateway.	
• Prefer to use the "Prefix/Suffix" feature in Gateway with label "Date Issued".	

Source

Element Name	Source

DC Definition	A related resource from which the described resource is derived.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Source (dc:source)
MARC Map in WorldCat	786 [08]
Repeatable	Yes

Best Practices

- Prefer use of free text description incl., Collection Name, Accession Number, Physical Dimensions for graphic materials and Repository information.
- Prefer "Original Format" or other text prefix to qualify value.

"Enter information about the original item before digitization as follows: genre of item: collection name, name of box, number of bin. Ex: 35 mm color slide: Larry Oglesby Collection, Morro Bay FT, bin #8" – Data Dictionary for Larry Oglesby Collection, LOC—Claremont Colleges Digital Library

Relation

Element Name	Relation
DC Definition	A related resource.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Relation (dc:relation)
MARC Map in WorldCat	787
Repeatable	Yes*

Best Practices

- Include sufficient information in the Relation element to enable users to identify, cite, and either locate or link to the related resource.
- *Some 'communities of practice' reference both the *Physical Collection* and the *Digital Collection*
- When applicable, use the more specific sub-elements

"The described resource is a physical or logical part of the referenced resource." – University of Wisconsin Digital Library Data Dictionary

Relation-Has Format Of

Element Name	Relation-Has Format Of
DC Definition	A related resource that is substantially the same as the pre-existing described resource, but in another format.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Has Format (dcterms:hasFormat)
MARC Map in WorldCat	776 08 $n
Repeatable	Yes

Best Practices

- The described resource is treated as the primary/pre-existing resource. For example, the

postcard, *"See Seattle" **postcard**, Alaska Yukon Pacific Exposition, 1909*, **Has Format Of** TIFF, scanned from original text, *"See Seattle" postcard **digital reproduction**, Alaska Yukon Pacific Exposition, 1909*, at 400 dpi. (See Relation-Is Format Of element).

Relation-Is Format Of

Element Name	Relation-Is Format Of
DC Definition	A related resource that is substantially the same as the described resource, but in another format.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Is Format Of (dcterms:isFormatOf)
MARC Map in WorldCat	776 08 $n
Repeatable	Not preferred
Best Practices	

- The described resource is treated as the secondary/supplement recourse. For example, the TIFF image, *"See Seattle" postcard **digital reproduction**, Alaska Yukon Pacific Exposition, 1909*, **Is Format Of** the original postcard, *"See Seattle" **postcard**, Alaska Yukon Pacific Exposition, 1909*, in 7 x 5 1/2 inch. (See Relation-Has Format Of element).

Relation-Has Part

Element Name	Relation-Has Part
DC Definition	A related resource that is included either physically or logically in the described resource.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Has Part (dcterms:hasPart)
MARC Map in WorldCat	774 08 $n
Repeatable	Yes
Best Practices	

"(For example)The described resource is an anthology that includes this article as well as other articles, each of which is described in another Relation [HasPart] element." - CDP Dublin Core Metadata Best Practices Version 2.1.

Relation-Is Part Of

Element Name	Relation-Is Part Of
DC Definition	A related resource in which the described resource is physically or logically included.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Is Part Of (dcterms:isPartOf)
MARC Map in WorldCat	773 0# $t
Repeatable	Not preferred
Best Practices	

- Used to state the collection to which this resource belongs. E.g., for Articles, this element indicates the host item (e.g., journal, series, etc.)

"The described resource is a physical or logical part of the referenced resource." – University of Wisconsin Digital Library Data Dictionary

Relation-Has Version

Element Name	Relation-Has Version
DC Definition	A related resource that is a version, edition, or adaptation of the described resource.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Has Version (dcterms:hasVersion)
MARC Map in WorldCat	775 08 $n
Repeatable	Yes
Best Practices	

- For example, Microsoft Office software **Has Version** Microsoft Office 97, Microsoft Office 2003, Microsoft Office 2010, etc.

Relation-Is Version Of

Element Name	Relation-Is Version Of
DC Definition	A related resource of which the described resource is a version, edition, or adaptation.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Is Version Of (dcterms:isVersionOf)
MARC Map in WorldCat	775 08 $n
Repeatable	Not preferred
Best Practices	

- For example, *The Lord of the rings* published by London: HarperCollins, in 2009 **Is Version Of** *The Lord of the rings* published by LONGMAN YORK PRESS in 1940s.

"Changes in version imply substantive changes in content rather than differences in format." - CDP Dublin Core Metadata Best Practices Version 2.1

Relation-Replaces

Element Name	Relation-Replaces
DC Definition	A related resource that is supplanted, displaced, or superseded by the described resource.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Replaces (dcterms:replaces)
MARC Map in WorldCat	780 00 $n
Repeatable	Yes

Best Practices
- For example, *Best Practices for CONTENTdm and other OAI-PMH compliant repositories 3.0* **Replaces** *Best Practices for CONTENTdm and other OAI-PMH compliant repositories 1.0*.

Relation-Is Replaced By

Element Name	Relation-Is Replaced By
DC Definition	A related resource that supplants, displaces, or supersedes the described resource.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Is Replaced By (dcterms:isReplacedBy)
MARC Map in WorldCat	785 00 $n
Repeatable	Yes

Best Practices
- For example, *Best Practices for CONTENTdm and other OAI-PMH compliant repositories 1.0* **Is Replaced By** *Best Practices for CONTENTdm and other OAI-PMH compliant repositories 3.0*.

Relation-Requires

Element Name	Relation-Requires
DC Definition	A related resource that is required by the described resource to support its function, delivery, or coherence.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Requires (dcterms:requires)
MARC Map in WorldCat	538
Repeatable	Yes

Best Practices
- This could be the technical information about an item. For example, a downloadable article **Requires** *Adobe Acrobat Reader, version 6.0*.

"When the resource being described requires the use of software, hardware, or other infrastructures that are external to the resource itself, record that information in the Relation [Requires] element. For example, if a Dublin Core record for the digitized version of a hand-written letter is delivered to the user as a PDF file, Adobe Acrobat Reader (which is external to the resource being described) is required to view that PDF file" – CDP Dublin Core Metadata Best Practices Version 2.1

Relation-Is Required By

Element Name	Relation-Is Required By
DC Definition	A related resource that requires the described resource to support its function, delivery, or coherence.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Is Required By (dcterms:isRequiredBy)
MARC Map in WorldCat	787 08 $n

Repeatable	Yes

Best Practices
- For example, the described resource is a life sciences dataset underline the scientific findings and **Is Required By** the paper, *Making Logistic Regression A Core Data Mining Tool With TR-IRLS*.

Coverage

Element Name	Coverage
DC Definition	The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Coverage (dc:coverage)
MARC Map in WorldCat	500
Repeatable	Yes

Best Practices
- A location, period of time, or jurisdiction of described resources. Spatial topic and spatial applicability may be a named place or a location specified by its geographic coordinates. Temporal topic may be a named period, date, or date range. A jurisdiction may be a named administrative entity or a geographic place to which the resource applies.
- For Spatial topic, prefer to use Coverage-Spatial element. For temporal topic, prefer to use Coverage-Temporal element.

"For artifacts or art objects, the spatial characteristics usually refer to the place where the artifact/object originated while the temporal characteristics refer to the date or time period during which the artifact/object was made." - CDP Dublin Core Metadata Best Practices Version 2.1

Coverage-Spatial

Element Name	Coverage-Spatial
DC Definition	Spatial characteristics of the resource.
Required	Recommended, as appropriate
Controlled Vocabulary	TGN, GNIS, LCNAF
Syntax Scheme	
DC Element Map	Spatial Coverage (dcterms:spatial)
MARC Map in WorldCat	522
Repeatable	Yes

Best Practices
- Prefer use of standard controlled vocabularies and name authority sources, such as Thesaurus of Geographic Names [TGN].
- Some 'communities of practice' reference geographic information system coordinates, such as those made available by Google Earth®

"Currently recommended by the "Collaborative Digitization Project Dublin Core Metatdata Best Practices" guide for use only 'in describing maps, globes, and cartographic resources or when place or time period cannot be adequately expressed using the Subject element.' Coverage spatial refers to the extent or scope of the content of the resource (e.g., place shown on a map or in a photograph, or

geographic locations that are the topic of a manuscript), not the place of publication or digitization." - Metadata Best Practices Guide, Western Michigan University Libraries

Coverage-Temporal

Element Name	Coverage-Temporal
DC Definition	Temporal characteristics of the resource.
Required	Recommended, as appropriate
Controlled Vocabulary	AAT, LCSH
Syntax Scheme	W3CDTF
DC Element Map	Temporal Coverage (dcterms:temporal)
MARC Map in WorldCat	648
Repeatable	Yes

Best Practices
- Use to describe the time period covered or represented by the resource, not the date when the resource was published. Temporal topic may be a named period, date, or date range.
- If using a named period, use a controlled vocabulary if possible such as Library of Congress Subjects (LCSH).
- Where appropriate, time periods can be date ranges in ISO 8601 W3C Date/Time Format standard.

"Usually a date or range of dates, but can be a named time period (e.g., Renaissance). Temporal coverage 'refers to the time period covered by the intellectual content of the resource (CDP Dublin Core Metadata Best Practices (CDPDCMBP)),' not the date of publication or digitization. It can refer to the time period shown in an image, the topic of a written manuscript, the time period covered in a series of diary entries, or, for art objects or artifacts, the date or time period of creation of the piece." - Metadata Best Practices Guide, Western Michigan University Libraries

Audience

Element Name	Audience
DC Definition	A class of entity for whom the resource is intended or useful.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Audience (dcterms:audience)
MARC Map in WorldCat	(521##$a enhancement recommended)
Repeatable	Yes

Best Practices
- Examples of Audience include students, women, charities, lecturers.

Provenance

Element Name	Provenance
DC Definition	A statement of any changes in ownership and custody of the resource since its creation that are significant for its authenticity, integrity, and interpretation.
Required	Recommended, as appropriate
Controlled Vocabulary	
Syntax Scheme	

DC Element Map	Provenance (dcterms:provenance)
MARC Map in WorldCat	(561##$a enhancement recommended)
Repeatable	Yes

Best Practices

- The statement may include a description of any changes successive custodians made to the resource.

"Provenance, from the French provenir, "to come from", refers to the chronology of the ownership or location of an historical object." - Oxford English Dictionary

References

[i] Members of the original CONTENTdm Metadata Working Group, Aug-Dec 2009

Sheila Bair	Western Michigan University	bair@wmich.edu
Dachun Bao	National Defense University	baod@ndu.edu
Amalia (Molly) Beisler	University of Nevada Reno	abeisler@unr.edu
Megan Bernal	Depaul University	MBERNAL2@depaul.edu
Laura Capell	University of Southern Mississippi	laura.capell@usm.edu
MingYu Chen	University of Houston	mchen15@uh.edu
Mei Ling Chow	Montclair University	chowm@mail.montclair.edu
Kevin Clair	Penn State University	kmc35@psulias.psu.edu
Lee Dotson	University of Central Florida	ddotson@mail.ucf.edu
Mario Einaudi	The Huntington Library	meinaudi@huntington.org
Allegra Gonzalez	Claremont Colleges Digital Library	Allegra_Gonzalez@cuc.claremont.edu
Deborah Green	University of Idaho	dgreen@uidaho.edu
Myung-Ja (MJ) Han	University of Illinois U-C	mhan3@illinois.edu
Rachel Howard	University of Louisville	rachel.howard@louisville.edu
Amanda A Hurford	Ball State University	aahurford@bsu.edu
Andrea Kappler	Evansville Vanderburgh Public Library	andreak@evpl.org
Deborah Keller	US Army	deborah.eb.keller@us.army.mil
Kate Kluttz	North Carolina State Library	kate.kluttz@ncdcr.gov
Lyn MacCorkle	University of Miami	LMaccork@miami.edu
Sandra McIntyre	Mountain West Digital Library	sandra.mcintyre@UTAH.EDU
Gail McMillan	Virginia Tech	gailmac@vt.edu
Ann Olszewski	Cleveland Public Library	ann.olszewski@cpl.org
Jennifer Palmentiero	SE NY Library Resources Council	jennifer@senylrc.org
Kitty Pittman	Oklahoma State Library	kpittman@oltn.odl.state.ok.us
Gayle Porter	Chicago State University	gporter@csu.edu
Gayle Spears	Atlanta University Center	gspears@auctr.edu
Jill Strass	St. Olaf University	strass@stolaf.edu
Glee M Willis	University of Nevada Reno	willis@unr.edu
Ling Wang	University of Illinois Chicago	lwang@uic.edu
Noelia Ramos	Map Library of Catalonia	noelia.ramos@icc.cat
Shilpa Rele	University of Miami	s.rele@miami.edu
Cheryl Walters	Utah State University	cheryl.walters@usu.edu
Trashinda Wright	Atlanta University Center	twright@auctr.edu
ZeeZee Zamin	Louisiana State University/LOUIS	zehra@lsu.edu

Joined 2012:

Natalie Bulick	Indiana State University	natalie.bulick@indstate.edu

[ii] Moving towards shareable metadata by Sarah L. Shreeves, Jenn Riley, and Liz Milewicz
First Monday, Volume 11, number 8 –7 (August 2006),
URL: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/issue/view/202

[iii] Han, Myung-Ja, Cho, Christine, Cole, Timothy W. and Jackson, Amy S. (2009) 'Metadata for Special Collections in CONTENTdm: How to Improve Interoperability of Unique Fields Through OAI-PMH', Journal of Library Metadata, 9: Issue3—4 , 213 — 238.

URL: http://dx.doi.org/10.1080/19386380903405124

Appendix A – Additional dcterms available

Identifier-Bibliographic Citation

Element Name	Identifier-Bibliographic Citation
DC Definition	A bibliographic reference for the resource.
Required	As Appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Bibliographic Citation (dcterms:bibliographicCitation)
MARC Map in WorldCat	(500 ##a enhancement recommended)
Repeatable	Yes
Best Practices	

- Used for Bibliographic Resource only.
- Recommended practice is to include sufficient bibliographic detail to identify the resource as unambiguously as possible.
- Prefer "Bibliographic citation" to qualify value.

Rights-License

Element Name	Rights-License
DC Definition	A legal document giving official permission to do something with the resource.
Required	As Appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	License (dcterms:license)
MARC Map in WorldCat	(540##$a enhancement recommended)
Repeatable	Yes
Best Practices	

Date-Modified

Element Name	Date-Modified
DC Definition	Date on which the resource was changed.
Required	As Appropriate
Controlled Vocabulary	
Syntax Scheme	W3CDTF
DC Element Map	Date Modified (dcterms:modified)
MARC Map in WorldCat	046 $j
Repeatable	Not preferred
Best Practices	

Relation-Conforms To

Element Name	Relation-Conforms To
DC Definition	An established standard to which the described resource conforms.
Required	As Appropriate

Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Conforms To (dcterms:conformsTo)
MARC Map in WorldCat	514 $e (Data Quality Note)
Repeatable	Yes
Best Practices	
• The standard is a basis for comparison; a reference point against which other things can be evaluated.	

Relation-References

Element Name	Relation-References
DC Definition	A related resource that is referenced, cited, or otherwise pointed to by the described resource.
Required	As Appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	References (dcterms:references)
MARC Map in WorldCat	787 08 $n
Repeatable	Yes
Best Practices	

Relation-Is Referenced By

Element Name	Relation-Is Referenced By
DC Definition	A related resource that references, cites, or otherwise points to the described resource.
Required	As Appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Is Referenced By (dcterms:isReferencedBy)
MARC Map in WorldCat	510 0#
Repeatable	Yes
Best Practices	

Audience-Education Level

Element Name	Audience-Education Level
DC Definition	A class of entity, defined in terms of progression through an educational or training context, for which the described resource is intended.
Required	As Appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Audience Education Level (dcterms:educationLevel)
MARC Map in WorldCat	(521##$a enhancement recommended)
Repeatable	Yes

Audience-Mediator

Element Name	Audience-Mediator
DC Definition	An entity that mediates access to the resource and for whom the resource is intended or useful.
Required	As Appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Mediator (dcterms:mediator)
MARC Map in WorldCat	
Repeatable	Yes
Best Practices	

- In an educational context, a mediator might be a parent, teacher, teaching assistant, or care-giver.

Instructional Method

Element Name	Instructional Method
DC Definition	A process, used to engender knowledge, attitudes and skills, that the described resource is designed to support.
Required	As Appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Instructional Method (dcterms:instructionalMethod)
MARC Map in WorldCat	
Repeatable	Yes
Best Practices	

- Instructional Method will typically include ways of presenting instructional materials or conducting instructional activities, patterns of learner-to-learner and learner-to-instructor interactions, and mechanisms by which group and individual levels of learning are measured. Instructional methods include all aspects of the instruction and learning processes from planning and implementation through evaluation and feedback.

Accrual Method

Element Name	Accrual Method
DC Definition	The method by which items are added to a collection.
Required	As Appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Accrual Method (dcterms:accrualMethod)
MARC Map in WorldCat	(541##$c enhancement recommended)
Repeatable	Yes
Best Practices	

- Used for Collection type of resource only.

Accrual Periodicity

Element Name	Accrual Periodicity
DC Definition	The frequency with which items are added to a collection. (Current Publication Frequency)
Required	As Appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Accrual Periodicity (dcterms:accrualPeriodicity)
MARC Map in WorldCat	(310##$a enhancement recommended)
Repeatable	Yes
Best Practices	

- Used for Collection type of resource only.

Accrual Policy

Element Name	Accrual Policy
DC Definition	The policy governing the addition of items to a collection.
Required	As Appropriate
Controlled Vocabulary	
Syntax Scheme	
DC Element Map	Accrual Policy (dcterms:accrualPolicy)
MARC Map in WorldCat	
Repeatable	Yes
Best Practices	

- Used for Collection type of resource only.
- A plan or course of action by an authority, intended to influence and determine decisions, actions, and other matters.

Appendix B: Moving Towards Marketing with Metadata

We have long recognized the need for effective marketing to increase discovery and delivery of digital collections. Enhancing descriptive metadata can move us in the right direction. Websites such as Flickr have adopted Web 2.0 social metadata standards such as tagging, in order to improve searchability for digital image material, and can leverage existing metadata to augment the user experience. There exists opportunity to further optimize descriptive metadata in otherwise well-aggregated digital collections. For example, there are many archival collections of historical material related to topics such as gold mining, railroad production, and other industries. The metadata used to describe these types of images can be quite literal and catalogers sometimes 'miss the point'-- failing to apply such key, albeit at times colloquial, descriptors as "boomtowns," "Gold Rush," or "Wild West."

While many controlled vocabularies are limited in their ability to incorporate this type of higher-level description, catalogers are encouraged to develop their own local controlled vocabularies based upon a convergence of subject terms (nouns, adjectives and verbs describing main topics) technical and style-based terms (unique image attributes such as image orientation, lens perspectives, and photographic techniques) and concept terms (ideas portrayed in an image). In WorldCat.org, the ability to

create/name lists of items and apply social tags to items allows a high level of flexibility in accessing and managing content. Thus, the further integration of digital content into WorldCat.org represents a unique opportunity for the special collections community to begin experimenting with these types of terminologies-focused workflow tasks to increase discovery.

Appendix C: Dates

Date type	DATE example
Known year-month-day	2001-10-19
Known year-month	2001-10
Known year	2001
One year or another	1892 or 1893
Circa year-month	circa 1843-02
Decade certain	1970s
Before a time period	before 1867
After a time period	after 1867

-Guidelines for Metadata Application in the Claremont Colleges Digital Library

About Dates in CONTENTdm:

1. CONTENTdm supports the "date" data type and is consistent with the ISO standard yyyy-mm-dd, yyyy-mm and yyyy. You must use the date data type in order to provide searchable dates in CONTENTdm. However, many CONTENTdm users also provide a date field using the text data type. The fields shown in the latter five examples above would need to be configured as "text".

2. To enter a range of years, use the following guidelines:

    a. CONTENTdm Project Client- Use the yyyy-yyyy standard. Upon saving your

    metadata, the CONTENTdm Project Client will break out every date in the range.

    b. CONTENTdm Web Add- Type every single year in the date range separated by

    semicolon-space.

-Metadata Implementation Guidelines for North Carolina Digital State Document

Appendix D: Metadata Schemas
The following are examples of CONTENTdm metadata schemas that represent the vetted work of the MWG:

	Field name	DC map	Data type	Large	Search	Hide	Required	Vocab		add field
1	Title	Title	Text	No	Yes	No	Yes	No	move to ⌄	edit | delete
2	Subject--Names	Subject	Text	No	Yes	No	No	Yes	move to ⌄	edit | delete
3	Subject-Places	Subject	Text	No	Yes	No	No	Yes	move to ⌄	edit | delete
4	Subject-Topics	Subject	Text	No	Yes	No	No	Yes	move to ⌄	edit | delete
5	Description	Description	Text	No	No	No	No	No	move to ⌄	edit | delete
6	Creator	Creator	Text	No	Yes	No	No	Yes	move to ⌄	edit | delete
7	Publisher	Publisher	Text	No	Yes	No	No	No	move to ⌄	edit | delete
8	Contributors	Contributors	Text	No	Yes	No	No	No	move to ⌄	edit | delete
9	Date Original	Date	Date	No	Yes	No	No	No	move to ⌄	edit | delete
10	Type	Type	Text	No	Yes	No	No	No	move to ⌄	edit | delete
11	Format	Format	Text	No	Yes	No	No	No	move to ⌄	edit | delete
12	Identifier	Identifier	Text	No	Yes	No	No	No	move to ⌄	edit | delete
13	Source	Source	Text	No	Yes	No	No	No	move to ⌄	edit | delete
14	Language	Language	Text	No	Yes	No	No	No	move to ⌄	edit | delete
15	Rights & Usage	Rights	Text	No	Yes	No	No	No	move to ⌄	edit | delete

For photographic collections (above) and archival collections (below)

	Field name	DC map	Data type	Large	Search	Hide	Required	Vocab	add field		
1	Title	Title	Text	No	Yes	No	Yes	No	move to ▾	edit	delete
2	Title-Alternative	Title-Alternative	Text	No	Yes	No	No	No	move to ▾	edit	delete
3	Creator	Creator	Text	No	Yes	No	No	Yes	move to ▾	edit	delete
4	Description	Description	Text	Yes	Yes	No	Yes	No	move to ▾	edit	delete
5	Abstract	Description-Abstract	Text	Yes	Yes	No	No	No	move to ▾	edit	delete
6	Table Of Contents	Description-Table Of Contents	Text	Yes	Yes	No	No	No	move to ▾	edit	delete
7	Transcription	Description	Full Text Search	Yes	Yes	No	No	No	move to ▾	edit	delete
8	Subject	Subject	Text	No	Yes	No	No	Yes	move to ▾	edit	delete
9	Publisher	Publisher	Text	No	Yes	No	No	No	move to ▾	edit	delete
10	Contributors	Contributors	Text	No	Yes	No	No	Yes	move to ▾	edit	delete
11	Date.Original	Date	Date	No	Yes	No	Yes	No	move to ▾	edit	delete
12	Date.Digital	None	Date	No	Yes	No	No	No	move to ▾	edit	delete
13	Type	Type	Text	No	Yes	No	Yes	Yes	move to ▾	edit	delete
14	Format	Format	Text	No	Yes	No	Yes	No	move to ▾	edit	delete
15	Identifier	Identifier	Text	No	Yes	No	Yes	No	move to ▾	edit	delete
16	Source	Source	Text	No	Yes	No	No	No	move to ▾	edit	delete
17	Language	Language	Text	No	Yes	No	No	No	move to ▾	edit	delete
18	Relation	Relation-Is Part Of	Text	No	Yes	No	No	No	move to ▾	edit	delete
19	Location	Coverage-Spatial	Text	No	Yes	No	No	Yes	move to ▾	edit	delete
20	Time Period	Coverage-Temporal	Text	No	Yes	No	No	Yes	move to ▾	edit	delete
21	Rights & Usage	Rights	Text	No	Yes	No	No	No	move to ▾	edit	delete
22	Audience	Audience	Text	No	Yes	No	No	No	move to ▾	edit	delete
23	Relation-References	Relation-References	Text	No	Yes	No	No	No	move to ▾	edit	delete
24	Digitization Specifications	None	Text	No	Yes	No	No	No	move to ▾	edit	delete

Appendix E: Compound Objects

Addendum on the treatment of compound objects with respect to OAI harvesting

Authors:

Geri Bunker Ingram, MLIS
OCLC Digital Collection
Services

Myung-Ja "MJ" Han
Metadata Librarian
Assistant Professor of Library
Administration
University of Illinois at
Urbana-Champaign

Sheila Bair, MLIS
Metadata & Cataloging
Librarian
Western Michigan University

Context:

During the drafting of the Best Practices Guide version 1.7, discussion arose among the Metadata Working Group concerning the special case of sharing metadata from CONTENTdm Compound Objects. Users may employ diverse strategies for sharing metadata, regardless of the material type or formats that are assembled as compound objects, and regardless of the OAI-PMH harvester that will be employed. A request was made to attach a statement to the guide explaining the implications of metadata schema definition and CONTENTdm field configuration when a collection containing Compound Objects is destined to be harvested.

CONTENTdm Definitions:

COMPOUND OBJECT –any two or more CONTENTdm items that are logically and structurally assembled together. Each compound object comprises:
- A metadata record describing the object itself, (known as *object-level metadata*).
- A metadata record (known as *page-level metadata*) for each of the composite pages or items that make up the compound object.

ITEM—a single digital file and its affiliated metadata. In cases where there is metadata only—e.g., an image has not yet been scanned, the metadata is known as a "metadata only item".

COMPOUND OBJECT CLASSES:
- Document—a series of related items
- Monograph—a series of items related in hierarchical fashion
- Post card—a series of exactly two items that may be displayed on one screen using the compound object viewer (by default labeled "front" and "back");
- Picture cube—a series of exactly six items (designed originally for scans of realia)

DOCUMENT DESCRIPTION (VIEW): One of several views of the compound object available from the 'compound object viewer'. The metadata that displays through this view is the *object-level* metadata.

PAGE DESCRIPTION (VIEW): One of several views of the compound object available from the 'compound object viewer'. The metadata that displays through this view is the *page-level* metadata.

Sharing metadata

With CONTENTdm, one can set a collection to be harvestable generally as long as the harvester is compliant, and one can also set a collection to be harvested by the Digital Collection Gateway specifically. With the former, CONTENTdm collection administrators can decide whether to enable the page-level metadata to be harvested. This is done in CONTENTdm Administration in the Server/Settings/OAI configuration function. With the Gateway, page-level metadata are never harvested, therefore the object-level metadata must be carefully considered. For other OAI harvesters, CONTENTdm collection administrators can decide whether and how fully to allow harvest of page-level metadata. Collection administrators should verify for every collection that the OAI configuration settings are correct for that particular collection.
The implications for discovery and delivery vary depending upon the type of object at hand, and how well the Compound object -level (*metadata of the object itself*) is represented. Collection administrators must determine whether the document description (object-level metadata) is enough for resource discovery/retrieval outside of the context of the native CONTENTdm environment. If a harvester provides direct links back to the object in its repository environment, (as in worldcat.org), and if the object-level metadata is extensive enough to allow discovery of the object, then end-users can link directly to the original collection and re-issue the specific search criteria to retrieve relevant objects with 'hits' highlighted on each page of each compound object across the collections on the server.

Example--Enhancing discovery of buried information

One of the CONTENTdm collections at Western Michigan University is a collection of Civil War diaries and letters assembled as compound objects. They employ the Library of Congress' "20 percent rule"[iii] for subject headings at the object level, except in cases of special information of interest to Civil War researchers. For instance, in all the diaries, subject headings at the object level contain the names of battles in which the diarist participated even though the description of the battle may comprise only a small percentage of the total text.

Special considerations for textual transcripts

The Document and Monograph classes of compound object in CONTENTdm are used mainly to handle text-rich objects. Searchable text transcripts are handled as metadata within a CONTENTdm schema. I.e., not only can every field of the metadata be made searchable, but above and beyond that, one field in each record may contain a searchable transcript of the text of the item. The *Full text search* field data type can be used for one field in each schema. In the case of a compound object, the object level metadata itself, and each of its item level metadata, may contain up to 128,000 characters in this Full text search field (often re-labeled "Transcript" in practice).CONTENTdm administrators decide whether to make this field harvestable or not, i.e., map the field to one of the DC elements.

Appendix F: Consortium issues

Addendum on considerations for consortia using OAI harvesting tools; adding value from the members' point of view

Authors:

Jason B. Lee
Metadata Coordinator,
WorldCat Digital Content
OCLC Digital Collection
Services

Lyn MacCorkle
Digital Project Development
& Repositories Librarian,
Digital Initiatives & Resources
University of Miami Libraries

Sandra McIntyre
Program Director, Mountain
West Digital Library

Gayle Porter
Special Formats Catalog
Librarian, University Library
Chicago State University

Taylor Surface
Senior Product Manager
OCLC Digital Collection
Services

Cheryl Walters
Head of Digital Initiatives,
Utah State University

Context:
A consortium is defined as an "agreement, combination, or group (as of companies) formed to undertake an enterprise beyond the resources of any one member." During the drafting of the Best Practices Guide ver. 1.7, discussion arose among Metadata Working Group members concerning digital production & syndication challenges from a consortial viewpoint. A task group was formed in order to identify these [primarily workflow-oriented] issues in order to set forth an additional suite of recommended guidelines and to propose and communicate some specific resolutions in the WorldCat Digital Gateway environment.

Considerations for Consortia:
We have identified several overlapping core considerations for institutional members of a consortium using OAI harvesting tools in order to contribute digital content to a central server (outside of the institution). These core considerations, which may affect workflows at both the institution- and consortium-levels, include but are not limited to, metadata practices, communication strategy, and coordination of tasks.

Note: In the CONTENTdm-specific scenarios we reference here, there are two distinctly different issues present:

1. One CONTENTdm license is owned by the Consortium and shared among institutions.
 2. One CONTENTdm license is owned as above, PLUS one or more CONTENTdm licenses are owned by member institutions.

Appendix G: Frequently Asked Questions regarding the Digital Collection Gateway
(see also http://www.oclc.org/gateway/)

1. Does the Digital Collection Gateway only allow a single registration (username and password) per server, and do all of the libraries in the consortium have to share login information?

Modifying or issuing Gateway license KEYS to accommodate multiple users, as well as multiple repositories, is the recommended workflow for consortia. A Gateway license key may allow up to 50 separate usernames for individual control of collections. The consortia should have some centralized control where all of the metadata is managed. This enables many user logins to the Gateway, facilitated by coordination with the repository system administrator to allow the metadata to be shared by OAI. Currently, any existing CONTENTdm user that is part of a consortium can send an e-mail request to contentdmsupport@oclc.org and request that their key be modified to 'allow xx number of users'. Once the change is implemented, **each** library consortia member would be able to create a separate Gateway registration @ https://worldcat.org/DigitalCollectionGateway/register.jsp [see Figure A below].

- Figure A: Digital Collection Gateway online registration page

2. Is there a way that multiple people can manage a repository in Digital Collection Gateway? It appears that when an admin *delegates* a collection to another person, he/she can no longer see or manage it.

In the Digital Collection Gateway interface, only one person can manage a repository at a time, but that means only that one person has control of the editing. Any user can go into the Manage Account tab and assign a collection to themselves or someone else. In other words, if 'Jason L.' is out on vacation for a while, then 'Taylor S.' can assign the "entire repository" collection to himself and manage the metadata map and sync schedule.

3. The set up and configuration for WorldCat Sync tasks is located in the Server tab in the CONTENTdm Web Administration area, which may only be accessible to staff at the institution-level. Therefore, who would need to perform the initial setup to enable each collection to be uploaded to the Digital Collection Gateway?

We recommend that staff write policies and procedures to clearly describe administrative tasks in OAI harvesting, such as initial registration/set-up & log-in information, record sync schedule, and selection of collections. These procedures need not be lengthy or laborious, but should be communicated and distributed to all institutions within the consortium.   Both the consortium staff and institutional staff need to coordinate their workflows to make sure that initial setup has been completed for each institution that wants to have their records added to the Gateway.

4. Would staff from both the consortia as well as the member library need to 'keep track' of which collections have been uploaded to the Gateway?

We recommend that consortia staff develop a reporting structure and make information standard and easily visible across stakeholder groups. Consortia staff should keep an up-to-date account of management of digital records through the OAI harvesting tool, so that members are aware of which records have been uploaded and to prevent duplication of effort. The Gateway now provides a monthly activity summary for an entire repository which details the number of records added, updated, and deleted on a collection by collection basis. Staff from both groups also need to be in agreement as to which collections are 'ready' to be uploaded to the Gateway as metadata is revised or updated in the repository, in preparation for a manual 'push' or automated regularly-set upload. Gateway users also now have the ability to block certain records from their collections from being loaded to WorldCat even if they are "published" in CONTENTdm. Staff from each institution who works with digital collections should understand and follow the consortia policies for managing their records.

5. What happens if digital records from a member library are harvested by the consortia, and then both the consortia and the member library upload those records to WorldCat?

Digital Collection Gateway, OCLC's self-service OAI harvesting tool, recently added an important identifier de-duplication enhancement for digital content uploaded to WorldCat. The Gateway will now verify that no other records in WorldCat contain the same item URL which will reduce the introduction of this type of duplication in WorldCat. Best practice calls for a consortium to identify a digital content syndication coordinator and task him/her with responsibility to coordinate contribution with an eye to quality and uniqueness, while minimizing duplication of effort among the membership.

6. In the consortial environment, what kind of metadata-specific practices do the partners need to agree upon?

Member libraries contributing digital content to a central server should agree on consistency in metadata-sharing practices by adopting a standard metadata style guide. Additionally, proprietary information such as rights, provenance, donor, etc., should be taken into consideration when determining what metadata is displayed locally, but not mapped for harvesting. For example, some consortia find it important to describe the process, equipment and specifications used to create the digital surrogate, although this information is often only useful within the local context. Mountain West Digital Library provides a non-Dublin Core field for this purpose (*Digitization Specifications*) which they adopted from the BCR/CDP DC Metadata Best Practices guide. Additionally, preservation data relating to archival master files are less useful in the aggregated environment, although a valuable best practice at the local level for migration purposes.

Consortia are also encouraged to develop a 'common field properties' schema that can be used flexibly for different types of materials such as theater programs, oral histories, and correspondence. Additionally, agreement and consistency (particularly in *level of granularity*) among the consortium on the intellectual content contained within digital collection records, particularly support the harvesting of shareable metadata related to:

- Subject & Genre information
- Geographic information
- Controlled vocabularies and name authorities
- *Required*, *Optional*, and *Recommended*, as well as *Searchable* designators
- Multiple field values vs. Repeating fields
- Display of qualifiers in the OAI environment
- *Original* Date vs. *Digitized* or *Published* Date
- Formatting conventions for Date, Language and other metadata fields